

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 June 2001 (07.06.2001)

PCT

(10) International Publication Number
WO 01/40312 A2

- (51) International Patent Classification⁸: C07K 1/00
- (21) International Application Number: PCT/GB00/04629
- (22) International Filing Date: 4 December 2000 (04.12.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
9928789.8 3 December 1999 (03.12.1999) GB
- (71) Applicant (*for all designated States except US*): DIVERSYS LIMITED [GB/GB]; 20 Park Crescent, London W1N 4AL (GB).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): HOLT, Lucy, Jessica [GB/GB]; 9b Anstey Way, Trumpington, Cambridge CB2 2JE (GB). DE WILDT, Rudolf, Maria, Theodora [GB/GB]; 27 Grange Road, Cambridge CB3 9AU (GB). TOMLINSON, Ian [GB/GB]; 13 Tunwells Lane, Great Shelford, Cambridge CB2 5LJ (GB).
- (74) Agents: MASCHIO, Antonio et al.; D Young & Co. 21 New Fetter Lane, London EC4A 1AD (GB).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— *Without international search report and to be republished upon receipt of that report.*
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



WO 01/40312 A2

(54) Title: NAIVE SCREENING METHOD

(57) Abstract: The invention describes a method for isolating, from a naïve polypeptide repertoire which has not been preselected with a specific target ligand, a polypeptide of interest capable of interacting with said specific target ligand, which method comprises direct screening of the naïve polypeptide repertoire with the target ligand in order to identify the polypeptide of interest.

Naïve Screening Method

The present invention relates to a method for screening polypeptide libraries in array form, without prior selection of the library using specific target ligands. In particular, the method relates to the use of enhanced immunoglobulin libraries in a direct screening protocol to isolate members of the libraries that bind a number of different expressed cDNAs.

Introduction

The antigen binding domain of an immunoglobulin comprises two separate regions: a heavy chain variable domain (V_H) and a light chain variable domain (V_L : which can be either V_K or V_λ). The antigen binding site itself is formed by six polypeptide loops: three from V_H domain (H1, H2 and H3) and three from V_L domain (L1, L2 and L3). A diverse primary repertoire of V genes that encode the V_H and V_L domains is produced by the combinatorial rearrangement of gene segments. The V_H gene is produced by the recombination of three gene segments, V_H , D and J_H . In humans, there are approximately 51 functional V_H segments (Cook and Tomlinson (1995) *Immunol Today*, 16: 237), 25 functional D segments (Corbett *et al.* (1997) *J. Mol. Biol.*, 268: 69) and 6 functional J_H segments (Ravetch *et al.* (1981) *Cell*, 27: 583), depending on the haplotype. The V_H segment encodes the region of the polypeptide chain which forms the first and second antigen binding loops of the V_H domain (H1 and H2), whilst the V_H , D and J_H segments combine to form the third antigen binding loop of the V_H domain (H3). The V_L gene is produced by the recombination of only two gene segments, V_L and J_L . In humans, there are approximately 40 functional V_K segments (Schäble and Zachau (1993) *Biol. Chem. Hoppe-Seyler*, 374: 1001), 31 functional V_λ segments (Williams *et al.* (1996) *J. Mol. Biol.*, 264: 220; Kawasaki *et al.* (1997) *Genome Res.*, 7: 250), 5 functional J_K segments (Hieter *et al.* (1982) *J. Biol. Chem.*, 257: 1516) and 4 functional J_λ segments (Vasicek and Leder (1990) *J. Exp. Med.*, 172: 609), depending on the haplotype. The V_L segment encodes the region of the polypeptide chain which forms the first and second antigen binding loops of the V_L domain (L1 and L2), whilst the V_L and J_L segments

combine to form the third antigen binding loop of the V_L domain (L3). Antibodies selected from this primary repertoire are believed to be sufficiently diverse to bind almost all antigens with at least moderate affinity. High affinity antibodies are produced by "affinity maturation" of the rearranged genes, in which point mutations are generated and
5 selected by the immune system on the basis of improved binding.

Analysis of the structures and sequences of antibodies has shown that five of the six antigen binding loops (H1, H2, L1, L2, L3) possess a limited number of main-chain conformations or canonical structures (Chothia and Lesk (1987) *J. Mol. Biol.*, **196**: 901;
10 Chothia *et al.* (1989) *Nature*, **342**: 877). The main-chain conformations are determined by (i) the length of the antigen binding loop, and (ii) particular residues, or types of residue, at certain key position in the antigen binding loop and the antibody framework. Analysis of the loop lengths and key residues has enabled us to predict the main-chain conformations of H1, H2, L1, L2 and L3 encoded by the majority of human antibody
15 sequences (Chothia *et al.* (1992) *J. Mol. Biol.*, **227**: 799; Tomlinson *et al.* (1995) *EMBO J.*, **14**: 4628; Williams *et al.* (1996) *J. Mol. Biol.*, **264**: 220). Although the H3 region is much more diverse in terms of sequence, length and structure (due to the use of D segments), it also forms a limited number of main-chain conformations for short loop lengths which depend on the length and the presence of particular residues, or types of
20 residue, at key positions in the loop and the antibody framework (Martin *et al.* (1996) *J. Mol. Biol.*, **263**: 800; Shirai *et al.* (1996) *FEBS Letters*, **399**: 1).

A similar analysis of side-chain diversity in human antibody sequences has enabled the separation of the pattern of sequence diversity in the primary repertoire from that created
25 by somatic hypermutation. It was found that the two patterns are complementary: diversity in the primary repertoire is focused at the centre of the antigen binding whereas somatic hypermutation spreads diversity to regions at the periphery that are highly conserved in the primary repertoire (Tomlinson *et al.* (1996) *J. Mol. Biol.*, **256**: 813; Ignatovich *et al.* (1997) *J. Mol. Biol.*, **268**: 69). This complementarity seems to have
30 evolved as an efficient strategy for searching sequence space, given the limited number B cells available for selection at any given time. Thus, antibodies are first selected from the primary repertoire based on diversity at the centre of the binding site. Somatic

hypermutation is then left to optimise residues at the periphery without disrupting favourable interactions established during the primary response.

The recent advent of phage-display technology (Smith (1985) *Science*, **228**: 1315; Scott and Smith (1990) *Science*, **249**: 386; McCafferty *et al.* (1990) *Nature*, **348**: 552) has enabled the *in vitro* selection of human antibodies against a wide range of target antigens from "single pot" libraries. These phage-antibody libraries can be grouped into two categories: natural libraries which use rearranged V genes harvested from human B cells (Marks *et al.* (1991) *J. Mol. Biol.*, **222**: 581; Vaughan *et al.* (1996) *Nature Biotech.*, **14**: 309) or synthetic libraries whereby germline V gene segments are 'rearranged' *in vitro* (Hoogenboom & Winter (1992) *J. Mol. Biol.*, **227**: 381; Nissim *et al.* (1994) *EMBO J.*, **13**: 692; Griffiths *et al.* (1994) *EMBO J.*, **13**: 3245; De Kruif *et al.* (1995) *J. Mol. Biol.*, **248**: 97) or where synthetic CDRs are incorporated into a single rearranged V gene (Barbas *et al.* (1992) *Proc. Natl. Acad. Sci. USA*, **89**: 4457). Although synthetic libraries help to overcome the inherent biases of the natural repertoire which can limit the effective size of phage libraries constructed from rearranged V genes, they require the use of long degenerate PCR primers which frequently introduce base-pair deletions into the assembled V genes. This high degree of randomisation may also lead to the creation of antibodies which are unable to fold correctly and are also therefore non-functional. Furthermore, antibodies selected from these libraries may be poorly expressed and, in many cases, will contain framework mutations that may effect the antibodies immunogenicity when used in human therapy.

Recently, in an extension of the synthetic library approach it has been suggested (WO97/08320, Morphosys) that human antibody frameworks can be pre-optimised by synthesising a set of 'master genes' that have consensus framework sequences and incorporate amino acid substitutions shown to improve folding and expression. Diversity in the CDRs is then incorporated using oligonucleotides. Since it is desirable to produce artificial human antibodies which will not be recognised as foreign by the human immune system, the use of consensus frameworks which, in most cases, do not correspond to any natural framework is a disadvantage of this approach. Furthermore, since it is likely that the CDR diversity will also have an effect on folding and/or expression, it is preferable to optimise the folding and/or expression (and remove any frame-shifts or stop codons) after

the V gene has been fully assembled. To this end, it is desirable to have a selection system which could eliminate non-functional or poorly folded/expressed members of the library before selection with the target antigen is carried out.

- 5 A solution to this, and other, problems is provided in our copending International patent application WO 99/20749 (MRC), the contents of which are incorporated herein by reference. The latter application describes a method by which a repertoire of polypeptides is preselected, according to functionality as determined by the ability to bind a generic ligand, and a subset of polypeptides obtained as a result of preselection is then employed
- 10 for further rounds of selection according to the ability to bind a target ligand. This approach permits the person skilled in the art to remove, from a chosen repertoire of polypeptides, those polypeptides which are non-functional, for example as a result of the introduction of frame-shift mutations, stop codons, folding mutants or expression mutants which would be or are incapable of binding to substantially any target ligand. Such non-
- 15 functional mutants are generated by the normal randomisation and variation procedures employed in the construction of polypeptide repertoires. At the same time, the invention permits the person skilled in the art to enrich a chosen repertoire of polypeptides for those polypeptides which are functional, well folded and highly expressed.
- 20 The generic ligands described for use with immunoglobulin repertoires in WO 99/20749 include superantigens. Superantigens are able to bind to functional immunoglobulin molecules, or subsets thereof comprising particular main-chain conformations, irrespective of target ligand specificity.
- 25 WO 99/02749 also describes libraries comprising a repertoire of polypeptides of the immunoglobulin superfamily, wherein the members of the repertoire have a known and/or single main-chain conformation. This has a number of advantages, including the elimination of non-binding immunoglobulins which are inappropriate to the target antigen as a result of an unsuitable main chain conformation.
- 30 WO 99/02749 describes improved libraries, but relies on display systems such as phage display in order to select desired polypeptides from the repertoires encoded by the libraries. Although phage display is very powerful, it is a laborious technique requiring

repeated rounds of phage growth, panning and infection, which often result in biases towards immuno-dominant epitopes and dominant proteins (in protein mixtures). Its use is necessitated by the size of antibody repertoire libraries, which in the prior art have not been selectable by other means.

5

Pre-selection using target ligands followed by screening for binding is common in the art, for example, by immunisation of a mouse, cloning the selected genes and performing colony lifts or by performing a round of phage or ribosome display and then screening the selected clones in an ELISA format. In the same way, target ligands can often be used to pre-select binders to different but related target ligands, presumably by binding common epitopes. Similarly, binders to a given target ligand can readily be diversified and screened against a related antigen without the use of phage display. Here we have devised a method for screening a polypeptide repertoire for specific members that interact with a given target ligand, whereby neither the target ligand nor any related molecule thereof has necessarily been in contact with any member of the polypeptide repertoire or any related polypeptide thereof.

10
15

Summary of the Invention

We have now developed a system which allows the selection of polypeptides from libraries in the absence of any pre-selection by exposure to target ligands without resorting to phage display or other display technologies for selection of polypeptides against the target ligands and without resorting to animal immunisation. The invention can be extended to the selection of polypeptides from any randomised polypeptide repertoire or cloned cDNA library.

20
25

In a first aspect of the present invention, therefore, there is provided a method for isolating from a polypeptide repertoire, which has not been preselected with a specific target ligand, a polypeptide of interest capable of interacting with the specific target ligand, which method comprises direct screening of the polypeptide repertoire with the target ligand in order to identify the specific polypeptide of interest.

30

Screening using the specific target ligand according to the invention is advantageously carried out in the absence of any genetic display package, such as phage display, polysome display or microencapsulation procedures.

5 Preferably, the polypeptide repertoire has been preselected with a generic ligand or ligands prior to screening against the target ligand. Such generic ligands are capable of binding only those molecules which are functionally expressed. We have found that this approach is one way in which the size of the library can be reduced sufficiently to allow direct screening of a target ligand against a polypeptide repertoire on an array, without
10 reducing the functionality of the library. Furthermore, it is also preferable that the target ligand that is chosen for screening is also functionally expressed. Such functional expression can be assayed using an appropriate generic ligand for the target ligand. More preferably, both the polypeptide repertoire and the target ligands are pre-selected by binding an appropriate generic ligand, before being screened against one another, thereby
15 reducing still further the number of polypeptides in the array that are required to isolate a specific binding pair.

The invention accordingly provides a method by which a repertoire of polypeptides is pre-selected, according to functionality as determined by the ability to bind the generic
20 ligand, and the subset of polypeptides obtained as a result of preselection is then employed for direct selection according to the ability to interact with the target ligand. In addition, the invention provides a method for checking the functionality of potential target ligands prior to screening by their ability to bind an appropriate generic ligand.

25 Pre-selection of the polypeptide repertoire removes those polypeptides which are non-functional, for example as a result of the introduction of frame-shift mutations, stop codons, folding mutants or expression mutants which would be or are incapable of binding to substantially any target ligand. Such non-functional mutants are generated by the normal randomisation and variation procedures employed in the construction of
30 polypeptide repertoires. At the same time, pre-selection enriches a chosen repertoire of polypeptides for those polypeptides which are functional, well folded and highly expressed. Pre-selection of the target ligands eliminates those that non-functional, for example as a result of the introduction of frame-shift mutations, stop codons, folding

mutants or expression mutants, which would be or are incapable of binding to substantially any member of the polypeptide repertoire.

Preferably, two or more subsets of polypeptides are obtained from a repertoire by the
5 method of the invention, for example, by prescreening the repertoire with two or more generic ligands, or by contacting the repertoire with the generic ligand(s) under different conditions. Advantageously, the subsets of polypeptides thus obtained are combined to form a further repertoire of polypeptides, which may be further screened by contacting
10 with target and/or generic ligands. More preferably, the target ligand is also pre-selected using with two or more generic ligand for the target ligand, or by contacting the target molecules with the generic ligand(s) under different conditions.

Preferably, the repertoire of polypeptides or the target molecules for use in the invention comprises polypeptides of the immunoglobulin superfamily, such as antibody
15 polypeptides or T-cell receptor polypeptides. Advantageously, the library may comprise individual immunoglobulin domains, such as the V_H or V_L domains of antibodies, or the V_β or V_α domains of T-cell receptors. In a preferred embodiment, therefore, repertoires of, for example, V_H and V_L polypeptides may be individually prescreened using a generic ligand and then combined to produce a functional repertoire comprising both V_H and V_L
20 polypeptides. Such a repertoire can then be screened with a target ligand in order to isolate polypeptides comprising both V_H and V_L domains and having the desired binding specificity. Alternatively, the screen can be reversed such that the target molecules comprise polypeptides of the immunoglobulin superfamily and the repertoire of polypeptides is, for example, an expressed cDNA library or a peptide library. In this case,
25 functional target molecules are first be selected by employing a screen using a generic ligand or ligands and then screened against the cDNA or peptide library. Such cDNA or peptide library can also be pre-screened by selection or screening with an appropriate generic ligand, such as an anti-tag antibody in the case of the cDNA library.

30 In an advantageous embodiment, the generic ligand selected for use with immunoglobulins is a superantigen. Superantigens are able to bind to functional immunoglobulin molecules, or subsets thereof comprising particular main-chain conformations, irrespective of target ligand specificity. Alternatively, generic ligands may

be selected from any ligand capable of binding to the general structure of the polypeptides which make up any given repertoire, such as antibodies themselves, metal ion matrices, organic compounds including proteins or peptides, and the like.

5 In a further aspect, the invention provides a library comprising members of the immunoglobulin superfamily, wherein the members of the repertoire have a known main-chain conformation. Such a library can be used to create an arrayed repertoire of immunoglobulin polypeptides for screening with a target ligand or as a source of target ligands for screening another polypeptide library.

10

Advantageously, the immunoglobulins have a single main-chain conformation. Although several main-chain conformations exist in nature, in practice it is possible to base an entire repertoire on a single main-chain conformation, without significant loss of functionality. This allows the library size to be reduced sufficiently to permit direct
15 screening of an arrayed repertoire.

Additionally, employing a known main-chain conformation maximises the ability of the repertoire to bind to generic ligands such as Protein A or Protein L, since the basic structure of the immunoglobulin fold is not expected to vary in such repertoires.

20

Preferably, a library based on an known and/or single main chain conformation may be pre-selected in accordance with the above aspect of the invention, in order to further refine the library by removing non-functional polypeptides from the repertoire.

25 In a preferred embodiment, the invention provides a naïve library or repertoire wherein the non-functional members have been completely or partially eliminated by binding with one or more generic ligands, said library or repertoire being suitable for direct screening with a specific target ligand or as a source of functional target ligands for screening another polypeptide library.

30

Preferably, the polypeptide repertoire according to the invention is designed for selection with both generic and target ligands. Advantageously, the target molecules are formatted for selection with at least one generic ligand and the repertoire of polypeptides.

- Repertoires according to the invention may be selected with the target ligand in a number of ways, some examples of which are set forth below. For example, the repertoires may be arrayed on a solid support and then screened using the target ligand. Alternatively the
- 5 repertoires may be screened in solution such that each member (or a small number of members) of the polypeptide repertoire is present in a separate well and the interaction with the target ligand occurs in the well containing the specific binding member of the polypeptide repertoire.
- 10 Repertoire arrays have many advantages, especially as comparative binding information may be obtained by comparing duplicate arrays, screened with different target ligands. Arrays may be created in a number of ways: for example, polypeptide repertoires may be arrayed by binding to a filter or other surface which may be coated with a generic ligand.
- 15 Alternatively, a polypeptide array of binding molecules, such as antibodies, may be arrayed directly by binding to the target ligand, which is itself bound to the filter. Only polypeptides which have specificity for the target ligand will be bound in this embodiment, meaning that the polypeptides can be visualised by binding with labelled generic ligand. Repertoires may be arrayed as such, in polypeptide form, or in the form of
- 20 an expression library of nucleic acids.

Preferably, polypeptide repertoires are screened using labelled target ligands. For example, if the polypeptide repertoire has been arrayed, the target ligand is advantageously labelled with an enzymatic label, such as HRP, a fluorescent label such as

25 FITC, or a radioactive label. If the polypeptide repertoire is in solution, the label is advantageously a moiety which can be isolated, such as one half of a binding pair, for example biotin or (strept)avidin. This facilitates the isolation of polypeptides which have interacted with (in this case bound to) the target ligand. Alternatively the bound target ligand can be detected using its own generic ligand, which are be advantageously labelled

30 with an enzymatic label, such as HRP, a fluorescent label such as FITC, or a radioactive label.

Advantageously, the repertoires of the invention (or the target molecules) comprise limited randomisation. A feature of side-chain diversity in the antigen binding site immunoglobulins such as human antibodies is a pronounced bias which favors certain amino acid residues. If the amino acid composition of the ten most diverse positions in
5 each of the V_H , V_K and V_L regions are summed, more than 76% of the side-chain diversity comes from only seven different residues, these being, serine (24%), tyrosine (14%), asparagine (11%), glycine (9%), alanine (7%), aspartate (6%) and threonine (6%). This bias towards hydrophilic residues and small residues which can provide main-chain flexibility probably reflects the evolution of surfaces which are predisposed to binding a
10 wide range of antigens and may help to explain the required promiscuity of antibodies in the primary repertoire.

Since it is preferable to mimic this distribution of amino acids, the invention provides a library wherein the distribution of amino acids at the positions to be varied mimics that
15 seen in the antigen binding site of antibodies. Such bias in the substitution of amino acids that permits selection of certain polypeptides (not just antibody polypeptides) against a range of target ligands is easily applied to any polypeptide repertoire according to the invention. There are various methods for biasing the amino acid distribution at the position to be varied (including the use of tri-nucleotide mutagenesis, WO97/08320, Morphosys, supra), of which the preferred method, due to ease of synthesis, is the use of
20 conventional degenerate codons. By comparing the amino acid profile encoded by all combinations of degenerate codons (with single, double, triple and quadruple degeneracy in equal ratios at each position) with the natural amino acid use it is possible to calculate the most representative codon. The codons (AGT)(AGC)T, (AGT)(AGC)C and
25 (AGT)(AGC)(CT) - that is, DVT, DVC and DVY, respectively using IUPAC nomenclature - are those closest to the desired amino acid profile: they encode 22% serine and 11% tyrosine, asparagine, glycine, alanine, aspartate, threonine and cysteine. Preferably, therefore, libraries are constructed using either the DVT, DVC or DVY codon at each of the diversified positions.

Brief Description of the Drawings

Figure 1

Unselected scFv D12 detects antigen M on array of expressed cDNAs from library hEX1
5 (Büssow *et al.*, 1998). Secondary detection is with recombinant protein L – HRP
(Affitech Ltd).

Figure 2

The sequence of the variable CDRs from each of the 4 naïve screened scFvs which
10 bound to an antigen on the array and the name of the antigen to which each scFv bound.
Each of these antigens is of unknown function.

Figure 3

A comparison between the detection sensitivity of a naïve screened scFv and a phage
15 selected scFv from the same scFv library. Naïve screened scFv D12 detects antigen M
and a phage selected anti ubiquitin scFv detects ubiquitin. The two scFvs detect their
target antigens with similar sensitivity. Secondary detection in both cases is with
recombinant protein L – HRP (Affitech Ltd).

20 Figure 4

A western blot with similar concentrations of antigen M and 10 other antigens was probed
with D12 scFv and only antigen M was detected. When antigen M was probed with a
scFv selected against BiPA4, then antigen M was not detected. Therefore D12 scFv is
specific for antigen M and antigen M is not cross reactive.

25

Figure 5

A Western blot showing the specificity of scFv 1 for antigen M and of scFvs 2 and 5 for
antigen D.

30 Figure 6

An ELISA showing the specificity of the light chain of clone 1 for antigen M and
showing the specificity of the heavy chains of clones 2 and 5 for antigen D.

Detailed Description of the Invention

Definitions

5 **Repertoire** A repertoire is a population of diverse variants, for example nucleic acid variants which differ in nucleotide sequence or polypeptide variants which differ in amino acid sequence. According to the present invention, a repertoire of polypeptides is preferably designed to possess a binding site for a generic ligand and a binding site for a target ligand. The binding sites may overlap, or be located in the same region of the molecule, but their specificities will differ.
10

Naïve Repertoire A repertoire is naïve if it has not previously been exposed to, or selected with, a specific target ligand. Typically, naïve repertoires of the prior art contain many redundant, non-functional or incomplete molecules. The repertoires according to the present invention comprise far fewer of such molecules, as a result of pre-selection with generic ligands, or deliberate construction using single or known main-chain conformations and limited randomisation. A naïve repertoire according to the present invention is one in which the repertoire members have not been exposed to the target ligand (either as a result of *in vitro* selection, for example, by phage, ribosome or puromycin display, by animal immunisation or by any other technique that leads to a preferential enrichment of the repertoire for repertoire members which specifically interact with the target ligand) prior to the screen according to the invention.
15
20

Functional As used herein, the term "functional" refers to a polypeptide of the repertoire or a target molecule which possesses either the native biological activity of its type, or any specific desired activity, for example as judged by its ability to fold correctly or be expressed as a full length product, without the presence of frame-shifts, stop codons and so on. A "functional" polypeptide of the repertoire will at least have the potential to interact with a given target molecule, whereas a non-functional polypeptide will have substantially no chance of interacting with the same target molecule. A "functional" target molecule will at least have the potential to interact with a member of the polypeptide repertoire, whereas a non-functional target molecule will have substantially no chance of interacting with a member of the polypeptide repertoire. The functionality of a given
25
30

polypeptide of the repertoire or of a given target molecule can be judged by its ability to bind its appropriate generic ligand, also defined below.

Generic ligand A generic ligand is a ligand that binds a substantial proportion of functional members in a given repertoire of polypeptides. Thus, the same generic ligand can bind many members of the repertoire regardless of their target ligand specificities (see below). In general, the presence of functional generic ligand binding site indicates that the repertoire member is expressed and folded correctly. Thus, binding of the generic ligand to its binding site provides a method for preselecting functional polypeptides from a repertoire of polypeptides. Target molecules may also have generic ligands that can be used to indicate the functionality of the target molecules.

Target ligand The target ligand is a molecule for which members of the polypeptide repertoire that have a specific activity are to be identified. Where the members of the polypeptide repertoire are antibody molecules, the target ligand may be an antigen and where the members of the repertoire are enzymes, the target ligand may be a substrate. Where the members of the polypeptide repertoire are expressed cDNAs, the target ligands may themselves be antibodies or some other polypeptide molecule. The specific activity of a target ligand - polypeptide repertoire member pair is dependent upon both the target ligand and the member of the repertoire being functional, as described above under *generic ligand*, and upon the precise specificity of the polypeptide repertoire member for the target ligand.

Activity The activity of a repertoire member polypeptide for the target ligand may be a binding activity, in which case the repertoire member binds the target molecule, an enzymatic activity, in which case it may catalyse a reaction involving the target ligand, or any other activity which is measurable.

Subset The subset is a part of the repertoire. In the terms of the present invention, it is often the case that only a subset of the repertoire is functional and therefore possesses a functional generic ligand binding site. Furthermore, it is also possible that only a fraction of the functional members of a repertoire (yet significantly more than would display

activity to a given target ligand) will bind the generic ligand. These subsets are able to be selected according to the invention.

Subsets of a library may be combined or pooled to produce novel repertoires which have been preselected according to desired criteria. Combined or pooled repertoires may be simple mixtures of the polypeptide members preselected by generic ligand binding, or may be manipulated to combine two polypeptide subsets. For example, V_H and V_L polypeptides may be individually prescreened, and subsequently combined at the genetic level onto single vectors such that they are expressed as combined V_H - V_L dimers, such as scFv.

Library The term library refers to a mixture of heterogeneous polypeptides or nucleic acids. The library is composed of members, which have a single polypeptide or nucleic acid sequence. To this extent, *library* is synonymous with *repertoire*. Where the *library* is a nucleic acid library, it encodes a *repertoire* of polypeptides. Sequence differences between library members are responsible for the diversity present in the library. The library may take the form of a simple mixture of polypeptides or nucleic acids, or may be in the form organisms or cells, for example bacteria, viruses, animal or plant cells and the like, transformed with a library of nucleic acids. Advantageously, nucleic acids are incorporated into expression vectors, in order to allow expression of the polypeptides encoded by the nucleic acids. In a preferred aspect, therefore, a library may take the form of a population of host organisms, each organism containing one or more copies of an expression vector containing a single member of the library in nucleic acid form which can be expressed to produce its corresponding polypeptide member. Thus, the population of host organisms has the potential to encode a large repertoire of genetically diverse polypeptide variants.

Immunoglobulin superfamily This refers to a family of polypeptides which retain the immunoglobulin fold characteristic of immunoglobulin (antibody) molecules, which contains two β sheets and, usually, a conserved disulphide bond. Members of the immunoglobulin superfamily are involved in many aspects of cellular and non-cellular interactions *in vivo*, including widespread roles in the immune system (for example, antibodies, T-cell receptor molecules and the like), involvement in cell adhesion (for

example the ICAM molecules) and intracellular signalling (for example, receptor molecules, such as the PDGF receptor). The present invention is applicable to all immunoglobulin superfamily molecules, since variation therein is achieved in similar ways. Preferably, the present invention relates to immunoglobulins (antibodies).

5

Main-chain conformation The main-chain conformation refers to the C α backbone trace of a structure in three-dimensions. When individual hypervariable loops of antibodies or TCR molecules are considered the main-chain conformation is synonymous with the canonical structure. As set forth in Chothia and Lesk (1987) *J. Mol. Biol.*, **196**: 901 and
10 Chothia *et al.* (1989) *Nature*, **342**: 877, antibodies display a limited number of canonical structures for five of their six hypervariable loops (H1, H2, L1, L2 and L3), despite considerable side-chain diversity in the loops themselves. The precise canonical structure exhibited depends on the length of the loop and the identity of certain key residues involved in its packing. The sixth loop (H3) is much more diverse in both length and
15 sequence and therefore only exhibits canonical structures for certain short loop lengths (Martin *et al.* (1996) *J. Mol. Biol.*, **263**: 800; Shirai *et al.* (1996) *FEBS Letters*, **399**: 1). In the present invention, all six loops will preferably have canonical structures and hence the main-chain conformation for the entire antibody molecule will be known.

20 **Antibody polypeptide** Antibodies are immunoglobulins that are produced by B cells and form a central part of the host immune defence system in vertebrates. An antibody polypeptide, as used herein, is a polypeptide which either is an antibody or is a part of an antibody, modified or unmodified. Thus, the term antibody polypeptide includes a heavy chain, a light chain, a heavy chain-light chain dimer, a Fab fragment, a F(ab')₂ fragment, a
25 Dab fragment, a light chain single domain, a heavy chain single domain, or an Fv fragment, including a single chain Fv (scFv). Methods for the construction of such antibody molecules are well known in the art.

Superantigen Superantigens are antigens, mostly in the form of toxins expressed in
30 bacteria, which interact with members of the immunoglobulin superfamily outside the conventional ligand binding sites for these molecules. Staphylococcal enterotoxins interact with T-cell receptors and have the effect of stimulating CD4⁺ T-cells. Superantigens for antibodies include the molecules Protein G that binds the IgG constant

region (Bjorck and Kronvall (1984) *J. Immunol.*, 133: 969; Reis *et al.* (1984) *J. Immunol.*, 132: 3091), Protein A that binds the IgG constant region and the V_H domain (Forsgren and Sjoquist (1966) *J. Immunol.*, 97: 822) and Protein L that binds the V_L domain (Bjorck (1988) *J. Immunol.*, 140: 1994). In the present invention, superantigens are used as
5 generic ligands for immunoglobulin superfamily members.

Preferred Embodiments of the Invention

The preparation of libraries and repertoires useful in the present invention is fully
10 described in WO 99/20749, the disclosure of which is incorporated herein by reference. WO 99/20749 describes how a library of immunoglobulins may be prepared and preselected using a generic ligand, and/or prepared using a single main-chain conformation. Libraries as described in WO 99/20749 can be expressed in host organisms, as described therein or according to techniques well known in the art, to
15 produce repertoires of polypeptides which are suitable for arraying and use in the present invention. Alternatively, polypeptides may be sequences and synthesised, *in situ* or otherwise, in arrays for use in the present invention.

Arraying of polypeptide repertoires

20

According to the present invention, polypeptides may be arrayed by any one of a variety of methods, depending upon whether the polypeptides are arrayed as such or expressed by arrayed cells. For example, polypeptides may be arrayed by robotic gridding using commercial technology as is commonly available in the art.

25

Arraying may be carried out by direct chemical synthesis of library members. One early method involves the synthesis of peptides on a set of pins or rods, such as described in WO84/03564. A similar method involving peptide synthesis on beads, which forms a peptide library in which each bead is an individual library member, is described in U.S.
30 Patent No. 4,631,211 and a related method is described in WO92/00091. A significant improvement of the bead-based methods involves tagging each bead with a unique identifier tag, such as an oligonucleotide, so as to facilitate identification of the amino

acid sequence of each library member. These improved bead-based methods are described in WO93/06121.

Another chemical synthesis method involves the synthesis of arrays of peptides (or peptidomimetics) on a surface in a manner that places each distinct library member (e.g., unique peptide sequence) at a discrete, predefined location in the array. The identity of each library member is determined by its spatial location in the array. The locations in the array where binding interactions between a predetermined molecule (e.g., a receptor) and reactive library members occur is determined, thereby identifying the sequences of the reactive library members on the basis of spatial location. These methods are described in U.S. Patent No. 5,143,854; WO90/15070 and WO92/10092; Fodor *et al.* (1991) *Science*, **251**: 767; Dower and Fodor (1991) *Ann. Rep. Med. Chem.*, **26**: 271.

(a) Arraying of cells

In a preferred aspect of the invention, arrays of nucleic acids may be prepared by arraying cells. Cells are advantageously arrayed by robotic picking, since robotic techniques allow the most precise and condensed gridding of cell colonies; however, any technique, including manual techniques, which is suitable for locating cells or colonies of cells at discrete locations on a support, may be used.

The gridding of cells may be regular, such that each colony is at a given distance from the next, or random. If colonies are spaced randomly, their density can be adjusted to statistically reduce or eliminate the probability of colonies overlapping on the chosen support.

Methods for arraying cell colonies are described in detail in US patent 5,326,691, the contents of which are incorporated herein by reference.

Screening Polypeptide Arrays in accordance with the Invention

Polypeptide arrays, once created, may be screened with a suitable target molecule or molecules in accordance with techniques known in the art. Advantageously, the target

ligand is labelled, or is itself detectable using a reagent which generates a detectable signal.

5 The invention may be configured according to a number of different criteria, which will dictate the use of different detection approaches. Preferred embodiments of the invention are set forth below, although other approaches will be apparent to those skilled in the art.

a. Solution approaches

10 The repertoire of polypeptides may be screened in solution. In this embodiment, the a subset of the repertoire may first be isolated using generic ligands and, where appropriate, multiple subsets may be combined prior to screening with the target ligand.

For solution screening each member (or a small number of members) of the polypeptide repertoire is present in a separate well of, for example, a 96, 384 or 1536 well plate. The
15 target ligand is then exposed to the repertoire, interactions with the target ligand occurring only in those wells that contain the specific binding member of the polypeptide repertoire. Such interactions can be detected by subsequent capture of the target ligand via a tag, such as a biotin molecule (in which case streptavidin is used to capture the target ligand) followed by the determination of the presence of the specific member of the repertoire,
20 perhaps by detection using a generic ligand.

b. Arraying approaches

An alternative set of embodiments includes methods which involve arraying the polypeptide repertoire on the target ligand, or binding the polypeptide repertoire to a solid
25 support, and exposing it to a labelled target ligand or a target ligand and, in a separate step, a label specific therefor.

In a first, preferred embodiment, the polypeptide repertoire is arrayed using a gridding process, such that the position of polypeptides in the array may be related to clones of a
30 library which encodes the repertoire.

More precisely, the target ligand may be immobilised on a solid surface, which is then exposed to the repertoire. Members of the repertoire that interact with the target ligand

can then be detected, perhaps using a labelled generic ligand directed against the functional members of the polypeptide repertoire.

Alternatively, the polypeptide repertoire may be arrayed directly onto the solid support
5 are via a generic ligand attached to a solid support. Whether the members of the polypeptide repertoire are fixed directly or indirectly to a solid support, the target ligands are advantageously labelled with a detectable label, to permit visualisation of the location of polypeptides on the array that interact with the target ligand. Alternatively, the binding of target ligands to the immobilised polypeptides can be detected by other means,
10 including for example surface plasmon resonance or the use of labelled generic ligands for the target molecule.

According to these embodiments, a repertoire may, if desired, be screened simultaneously against a number of different target ligands to identify members of the polypeptide
15 repertoire that interact with a desired combination of target ligands.

Robotic arraying is well-known in the art, and machines are available from companies such as Genetix, Genetic MicroSystems and BioRobotics which are capable of arraying at high speed with great accuracy over small or large surfaces. Such machines are capable
20 of spotting purified protein, supernatant or cells onto porous or non-porous surfaces, such that they can subsequently be fixed thereto if necessary to produce stable arrays. As described above, arrays can be replicated if required, to allow simultaneous screening with multiple target ligands. Cell-based arrays can be lysed to release polypeptides *in situ*, and/or expressed polypeptides can be fixed to the solid support according to known
25 procedures.

Use of polypeptides selected according to the invention

Polypeptides selected according to the method of the present invention may be employed
30 in substantially any process. Where the polypeptides are antibody polypeptides, they may be used in any process which involves ligand-polypeptide binding, including *in vivo* therapeutic and prophylactic applications, *in vitro* and *in vivo* diagnostic applications, *in vitro* assay and reagent applications, and the like. For example, in the case of antibodies,

antibody molecules may be used in antibody based assay techniques, such as ELISA techniques, according to methods known to those skilled in the art.

As alluded to above, the molecules selected according to the invention are of use in
5 diagnostic, prophylactic and therapeutic procedures. For example, enzyme variants generated and selected by these methods may be assayed for activity, either *in vitro* or *in vivo* using techniques well known in the art, by which they are incubated with candidate substrate molecules and the conversion of substrate to product is analysed. Selected cell-
10 surface receptors or adhesion molecules are be expressed in cultured cells which are then tested for their ability to respond to biochemical stimuli or for their affinity with other cell types that express cell-surface molecules to which the undiversified adhesion molecule are be expected to bind, respectively. Antibody polypeptides selected according to the invention are of use diagnostically in Western analysis and *in situ* protein detection by standard immunohistochemical procedures; for use in these applications, the antibodies of
15 a selected repertoire may be labelled in accordance with techniques known to the art. In addition, such antibody polypeptides may be used preparatively in affinity chromatography procedures, when complexed to a chromatographic support, such as a resin. All such techniques are well known to one of skill in the art.

20 Therapeutic and prophylactic uses of proteins prepared according to the invention involve the administration of polypeptides selected according to the invention to a recipient mammal, such as a human. Of particular use in this regard are antibodies, other receptors (including, but not limited to T-cell receptors) and in the case in which an antibody or receptor was used as either a generic or target ligand, proteins which bind to them.

25 Substantially pure antibodies or binding proteins thereof of at least 90 to 95% homogeneity are preferred for administration to a mammal, and 98 to 99% or more homogeneity is most preferred for pharmaceutical uses, especially when the mammal is a human. Once purified, partially or to homogeneity as desired, the selected polypeptides
30 may be used diagnostically or therapeutically (including extracorporeally) or in developing and performing assay procedures, immunofluorescent staining and the like (Lefkovite and Pernis, (1979 and 1981) Immunological Methods, Volumes I and II, Academic Press, NY).

The selected antibodies or binding proteins thereof of the present invention will typically find use in preventing, suppressing or treating inflammatory states, allergic hypersensitivity, cancer, bacterial or viral infection, and autoimmune disorders (which
5 include, but are not limited to, Type I diabetes, multiple sclerosis, rheumatoid arthritis, systemic lupus erythematosus, Crohn's disease and myasthenia gravis).

In the instant application, the term "prevention" involves administration of the protective composition prior to the induction of the disease. "Suppression" refers to administration
10 of the composition after an inductive event, but prior to the clinical appearance of the disease. "Treatment" involves administration of the protective composition after disease symptoms become manifest.

Animal model systems which can be used to screen the effectiveness of the antibodies or
15 binding proteins thereof in protecting against or treating the disease are available. Methods for the testing of systemic lupus erythematosus (SLE) in susceptible mice are known in the art (Knight *et al.* (1978) *J. Exp. Med.*, **147**: 1653; Reinersten *et al.* (1978) *New Eng. J. Med.*, **299**: 515). Myasthenia Gravis (MG) is tested in SJL/J female mice by inducing the disease with soluble AchR protein from another species (Lindstrom *et al.*
20 (1988) *Adv. Immunol.*, **42**: 233). Arthritis is induced in a susceptible strain of mice by injection of Type II collagen (Stuart *et al.* (1984) *Ann. Rev. Immunol.*, **42**: 233). A model by which adjuvant arthritis is induced in susceptible rats by injection of mycobacterial heat shock protein has been described (Van Eden *et al.* (1988) *Nature*, **331**: 171). Thyroiditis is induced in mice by administration of thyroglobulin as described (Maron *et al.*
25 *al.* (1980) *J. Exp. Med.*, **152**: 1115). Insulin dependent diabetes mellitus (IDDM) occurs naturally or can be induced in certain strains of mice such as those described by Kanasawa *et al.* (1984) *Diabetologia*, **27**: 113. EAE in mouse and rat serves as a model for MS in human. In this model, the demyelinating disease is induced by administration of myelin basic protein (see Paterson (1986) *Textbook of Immunopathology*, Mischer *et al.*, eds., Grune and Stratton, New York, pp. 179-213; McFarlin *et al.* (1973) *Science*,
30 **179**: 478; and Satoh *et al.* (1987) *J. Immunol.*, **138**: 179).

The selected antibodies, receptors (including, but not limited to T-cell receptors) or binding proteins thereof of the present invention may also be used in combination with other antibodies, particularly monoclonal antibodies (MAbs) reactive with other markers on human cells responsible for the diseases. For example, suitable T-cell markers can
5 include those grouped into the so-called "Clusters of Differentiation," as named by the First International Leukocyte Differentiation Workshop (Bernhard *et al.* (1984) *Leukocyte Typing*, Springer Verlag, NY).

Generally, the present selected antibodies, receptors or binding proteins will be utilised in
10 purified form together with pharmacologically appropriate carriers. Typically, these carriers include aqueous or alcoholic/aqueous solutions, emulsions or suspensions, any including saline and/or buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride and lactated Ringer's. Suitable physiologically-acceptable adjuvants, if necessary to keep a polypeptide complex in
15 suspension, may be chosen from thickeners such as carboxymethylcellulose, polyvinylpyrrolidone, gelatin and alginates.

Intravenous vehicles include fluid and nutrient replenishers and electrolyte replenishers, such as those based on Ringer's dextrose. Preservatives and other additives, such as
20 antimicrobials, antioxidants, chelating agents and inert gases, may also be present (Mack (1982) *Remington's Pharmaceutical Sciences*, 16th Edition).

The selected polypeptides of the present invention may be used as separately administered compositions or in conjunction with other agents. These can include various
25 immunotherapeutic drugs, such as cyclosporine, methotrexate, adriamycin or cisplatinum, and immunotoxins. Pharmaceutical compositions can include "cocktails" of various cytotoxic or other agents in conjunction with the selected antibodies, receptors or binding proteins thereof of the present invention, or even combinations of selected polypeptides according to the present invention having different specificities, such as polypeptides
30 selected using different target ligands, whether or not they are pooled prior to administration.

The route of administration of pharmaceutical compositions according to the invention may be any of those commonly known to those of ordinary skill in the art. For therapy, including without limitation immunotherapy, the selected antibodies, receptors or binding proteins thereof of the invention can be administered to any patient in accordance with
5 standard techniques. The administration can be by any appropriate mode, including parenterally, intravenously, intramuscularly, intraperitoneally, transdermally, *via* the pulmonary route, or also, appropriately, by direct infusion with a catheter. The dosage and frequency of administration will depend on the age, sex and condition of the patient, concurrent administration of other drugs, counterindications and other parameters to be
10 taken into account by the clinician.

The selected polypeptides of this invention can be lyophilised for storage and reconstituted in a suitable carrier prior to use. This technique has been shown to be effective with conventional immunoglobulins and art-known lyophilisation and
15 reconstitution techniques can be employed. It will be appreciated by those skilled in the art that lyophilisation and reconstitution can lead to varying degrees of antibody activity loss (e.g. with conventional immunoglobulins, IgM antibodies tend to have greater activity loss than IgG antibodies) and that use levels may have to be adjusted upward to compensate.

20 The compositions containing the present selected polypeptides or a cocktail thereof can be administered for prophylactic and/or therapeutic treatments. In certain therapeutic applications, an adequate amount to accomplish at least partial inhibition, suppression, modulation, killing, or some other measurable parameter, of a population of selected cells
25 is defined as a "therapeutically-effective dose". Amounts needed to achieve this dosage will depend upon the severity of the disease and the general state of the patient's own immune system, but generally range from 0.005 to 5.0 mg of selected antibody, receptor (e.g. a T-cell receptor) or binding protein thereof *per* kilogram of body weight, with doses of 0.05 to 2.0 mg/kg/dose being more commonly used. For prophylactic applications,
30 compositions containing the present selected polypeptides or cocktails thereof may also be administered in similar or slightly lower dosages.

A composition containing a selected polypeptide according to the present invention may be utilised in prophylactic and therapeutic settings to aid in the alteration, inactivation, killing or removal of a select target cell population in a mammal. In addition, the selected repertoires of polypeptides described herein may be used extracorporeally or *in vitro* selectively to kill, deplete or otherwise effectively remove a target cell population from a heterogeneous collection of cells. Blood from a mammal may be combined extracorporeally with the selected antibodies, cell-surface receptors or binding proteins thereof whereby the undesired cells are killed or otherwise removed from the blood for return to the mammal in accordance with standard techniques.

The invention is further described, for the purpose of illustration only, in the following examples.

Example 1

Screening of an expressed cDNA array using antibodies as target ligands

Target ligands for the screen were chosen from Tomlinson libraries I and J, which are based on a single human framework for VH (V3-23/DP47 and JH4b) and for VK (O12/O2/DPK9 and JK1). In this scFv library, side chain diversity is incorporated by PCR into CDRs 2 and 3 of both the heavy and light chains. The canonical structure that is encoded by this framework (VH: 1-3, VK: 2-1-1) is the most common in the human antibody repertoire. The antibody genes are cloned in vector pIT2 for IPTG inducible expression of the scFvs, which are directed to the bacterial periplasm by the pelB leader sequence. During scFv expression, scFvs diffuse out of the periplasm into the supernatant. The scFv libraries have been preselected for VH binding to bacterial superantigen A and VK binding to bacterial superantigen L. 96 ScFvs were grown overnight on 100µg/ml ampicillin, 1% glucose TYE media, then picked and then grown overnight in 100µg/ml ampicillin, 1% glucose 2xTY liquid culture. A small inoculum (about 2 µl) was transferred to a fresh 96 well plate (with a 96 well transfer device) and clones were regrown at 37 °C for 3h to OD 0.9. IPTG was added and the clones were induced overnight. Bacterial cells were pelleted by centrifugation and the supernatant was tested in ELISA for scFv expression level. The twelve highest expressing clones from a plate of 96 were chosen using a generic ligand sandwich assay. In this assay,

Protein L is bound to a plate, supernatants from the individual clones are incubated on the Protein L and then bound scFvs are detected using Protein A-HRP. Those clones which are functional and well expressed give a strong signal in this ELISA, whereas those clones which are non-functional or poorly expressed give a low signal or no signal at all.

- 5 The twelve scFvs chosen for high expression were individually repicked, regrown and induced to produce 100mls of filter sterilised scFv supernatant.

Two duplicate PVDF membranes each with 27,000 double spotted clones from cDNA expression library hEX1 (Bussow et al. 1998) were soaked in absolute ethanol for 20
10 minutes. The membranes were then each washed with agitation for 5 minutes in 1 litre of PBS/0.5% Triton X-100/0.05% Tween 20 (PBS-T-T). Each membrane was then rinsed twice briefly in 1 litre of PBS and washed for 5 minutes with agitation in a further 1 litre of PBS. Each membrane was then blocked for 45 minutes with agitation in 1 litre of 3 % MPBS, followed by 1 hour incubation with agitation in 6 x 10 ml scFv supernatant and 60
15 ml 2 x PBS, 6 % marvel. One of the two membranes was incubated with the 6 scFvs from library I and the other was incubated with the 6 scFvs from library J. After the 1 hour incubation with scFv, each membrane was washed with agitation for 5 minutes in 1 litre of PBS-T-T, followed by two rinses each in 1 litre of PBS and one five minute wash with agitation in 1 litre of PBS. Each membrane was then incubated for 40 minutes in 150
20 ml of a 1 in 5000 dilution of protein L-HRP (Affitech) in 3 % MPBS. Each membrane was then washed twice for 5 minutes with agitation in 1 litre of PBS-T-T and twice with agitation for 5 minutes in 1 litre of PBS. Each filter was then developed in a total volume of 60 ml of ECL reagents (Amersham Life Science) and exposed to photographic film (Kodak). As a control, to show that scFvs bind their target antigen in this system, a scFv
25 that had previously been identified by phage selection as a binder to ubiquitin was used to probe a membrane with library hEX1 arrayed. Expressed cDNA clones that were identified as positives were sequenced. The two clones sequenced to date were both found to be ubiquitin.

- 30 In total four different antibody-antigen interactions were identified without the need for phage selection against the target ligand and these were confirmed by Western blotting.

Antigens that were bound by one of the 12 scFvs were identified on the photographic film when both of their double spotted locations gave a positive signal (Fig 1).

In total four different antibody – antigen interactions were identified without the need for phage selection against the target ligand (Fig. 2 and Table 1). These were confirmed by Western blotting.

One of the four antibody antigen pairings (D12 and antigen M) was tested further, for specificity and sensitivity of detection of the naïve screened scFv for its target antigen. The sensitivity of detection of scFv D12 for its target, antigen M, was tested on a Western blot (Fig. 3). The sensitivity of detection of D12 for antigen M was compared to the sensitivity of detection of an anti ubiquitin scFv for its target antigen, ubiquitin. The phage selected and naïve screened scFvs detected their target antigens with similar sensitivity. The specificity of scFv for its target antigen, D12, was also tested on a Western blot (Fig. 4). Similar concentrations of antigen M and 10 other antigens were probed with D12 scFv and only antigen M was detected. When antigen M was probed with a scFv selected against BiPA4, antigen M was not detected. Therefore scFv D12 showed specific detection of antigen M and antigen M was not shown to be cross reactive

Table 1

Amino acid sequence of each of the antigens identified in the direct screen

Amino acid sequence of C

MRGSHHHHHGSYLGDTIESSTHASAHASAHASGPADKRAVRACGQGKE
ELVDPQIAPSSPPLPFCLEVPGSVFTDIWGKKRSRKLPRTELSRKQQQHHAKTLQR
KRRKAKVQRAKEGGKTPVWRI*

Amino acid sequence of O

MRGSHHHHHGSYLGDTIESSTHASAERVVGAKLRKGECGEGPGVVGPG
VWDRRVLTADFLGFALSSPEKTPGRAAHAGER*

Amino acid sequence of B

HHHHHHGSYLGDTHESSTHASGTCSQSFPCCCLCCAACPVCWAPPSPASAL
PRPPRLSGSSCSRHPGCAPGPSGCSACAQVPSGGRASGLADPAPVAVAAARCTP
TETKLLIS

5

Amino acid sequence of M

HHHHHHGSYLGDTHESSTHASGEKESSAVPARSLSSSPQAQPPRPAELSDE
EVAELFQRLAETQQEKWMLEEKVKHLEVSSASMAEDLCRKSIIETVYVMDSRID
VSVAAGHTDRSGLGSVLRDLVKPGDENLREMNNKKLQNMLEEQLTKNMHLHKD
MEVLSQEIVRLSKECVGPPDPDLEPGETS*

10

Example 2**Screening of arrayed scFv library using expressed cDNAs as target ligands****15 Creation of antibody array**

A naïve array of unselected antibodies was created. The method used to create the antibody array was as described in copending International patent application WO 99/20749 (MRC) and in de Wildt *et al.*, 2000, except that unselected antibody clones that have not been exposed to any target antigen (not those from the first round of phage selection) were used. Also, the antibody libraries used were cloned into the pAB1 vector, which does not have the phage P3 protein. These libraries were used to eliminate the possibility that any antibody-antigen interactions identified by naïve screening were as a result of contamination from phage selected antibodies. Briefly, either 18,432 single unselected clones from Tomlinson libraries I and J were picked into liquid culture in 384-well plates and grown overnight at 37°C or pools of clones were transferred to each well of the 384 well plates. (This was done by growing bacterial cells at a density of approximately 10^6 per plate (genetix Q-tray) and then stamping a 384-pin head into the plate and transferring bacteria on the end of the pins to a 384-well plate. This resulted in the transfer of multiple bacteria to each well of the 384 well plate and therefore allowed more clones to be screened for the same density of gridded bacteria on the membrane.) Bacteria from overnight culture were gridded in duplicate onto two nitrocellulose membranes that had been pre-blocked in 4% Marvel and was each covering an agar plate (Genetix Q-tray 100 µg/ml Ampicillin, 1% glucose 2xTy). The gridded plates were

20

25

30

grown overnight at 37°C . Meanwhile, target ligands for the naïve screen were chosen from the human brain cDNA library hEX1 (Bussow *et al.*, 1998) based on their ability to produce a well expressed protein product. Protein products were divided into two pools and coated overnight onto two nitrocellulose membranes (22cm x22cm) at approximately 5 10 µg/ml in 100ml PBS. The antigen membranes were washed 2 times in PBS and then blocked in 4% Marvel-PBS for 1 hour, washed 2 x in PBS, rinsed 1 x in 2xTy and then transferred to large square plates (Genetix Q-tray containing 250 ml 100 µg/ml Amp, 1mM isopropyl β-D-galactoside (IPTG) 2xTy. The first membrane containing the grown bacteria was then transferred onto the second (antigen) membrane in each case. The 10 plates were incubated for 3h to induce expression of the scFvs and facilitate their binding to the target ligand.

Probing of membranes

The top (bacterial) membranes were discarded and the bottom membranes were washed 2 x in PBS (0.05% tween-20) (PBST) and then blocked in 4% MPBS for 30 mins, followed 15 by two washes with PBST. For detection of bound scFvs, the membranes were incubated for 1 hour with 1/2000 protein L-HRP in MPBS for 1 hour, washed two times with PBST and then developed with ECL reagent followed by exposure to photographic film (Kodak).

Identification of binding antibodies

20 Two membranes coated with different target antigens were probed with the same collection of antibodies. Any scFvs showing a significant difference in binding strength on the two different antigens were picked and streaked onto agar plates. Single colonies were picked from the re-streaked plates and these putative positive clones were re-screened against individual proteins in a scaled-down array format. (In the case of the 25 pooled scFvs, 12-16 colonies were rescreened from each putative positive.) Any leads were identified, grown up in large volumes of 2xTy and induced to express soluble scFv. These scFvs were then tested for binding to antigen in both ELISA and Western format.

Results

Three scFvs were identified that showed specificity for their target antigen. ScFv 1 is 30 specific for antigen M and scFvs 2 and 5 are specific for antigen D. All three scFvs were shown to be in vector pAB1 (and therefore not selected contaminants on the naïve screen). All three scFvs bind target antigen specifically in both ELISA and Western (Figure 1) formats. Interestingly, all three scFvs was found to have binding specificity

residing principally in either the heavy or the light chain. The light chain of scFv 1 binds antigen M when paired with the dummy heavy chain and the heavy chains of scFvs 2 and 5 bind antigen D when paired with dummy light chains (Figure 2). Clones 1, 2 and 5 were sequenced and the sequences are given in Table 2 below

5

Table 2

Clone	Target antigen	CDR H2	CDR H3	CDR L2	CDR L3
1	M	TIPAVGRBTTYADSVKG	SQLPFDY	HASTLQS	QQQPLPPPT
2	D	TIFSKGSPTSADSVKG	TRVAFDY	RASRLQS	QQGEGFPDT
5	D	MITARGSPTSADSVKG	ITPGFDY	FASSLQS	QQARRAPPT

All publications mentioned in the above specification, and the references cited in said publications, are herein incorporated by reference.

10

Various modifications and variations of the described methods and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

15

Claims

1. A method for isolating, from a naïve polypeptide repertoire which has not been preselected with a specific target ligand, a polypeptide of interest capable of interacting
5 with said specific target ligand, which method comprises direct screening of the naïve polypeptide repertoire with the target ligand in order to identify the polypeptide of interest.
2. A method according to claim 1, wherein the naïve polypeptide repertoire is
10 preselected with a generic ligand, which is capable of binding those molecules in the polypeptide repertoire which are functionally expressed.
3. A method according to claims 1 or 2 wherein the generic ligand binds a subset of the repertoire of polypeptides.
15
4. A method according to claim 3 wherein two or more subsets are selected from the repertoire of polypeptides.
5. A method according to claim 4 wherein the selection is performed with two or
20 more generic ligands.
6. A method according to claims 4 or 5 wherein the two or more subsets are combined after selection to produce a further repertoire of polypeptides.
- 25 7. A method according to any preceding claim wherein two or more repertoires of polypeptides are contacted with generic ligands and the subsets of polypeptides thereby obtained are then combined.
8. A method according to claim 1 wherein a functionally expressed target ligand is
30 chosen from a group of target ligands based on its ability to bind one or more generic ligands.

9. A method according to any preceding claim, wherein the polypeptides of the repertoire or the target ligands are of the immunoglobulin superfamily.
10. A method according to claim 8, wherein the polypeptides of the repertoire or the target ligands are antibody or T-cell receptor polypeptides.
11. A method according to claim 9, wherein the polypeptides of the repertoire or the target ligands are V_H or V_β domains.
12. A method according to claim 9, wherein the polypeptides of the repertoire or the target ligands are V_L or V_α domains.
13. A method wherein a repertoire of polypeptides according to claim 11 and a repertoire of polypeptides according to claim 12 are contacted with generic ligands and the subsets thereby obtained are then combined.
14. A method according to any preceding claim wherein the generic ligand is selected from the group consisting of a matrix of metallic ions, an organic compound, a protein, a peptide, a monoclonal antibody, a polyclonal antibody population, and a superantigen.
15. A method for detecting, immobilising, purifying or immunoprecipitating one or more members of the repertoire of polypeptides previously selected according to any one of claims 1 to 14, comprising binding the members to the generic ligand.
17. A method according to any preceding claim, wherein the nucleic acid molecules encoding the repertoire of polypeptides are arranged at random to create an array.
18. A method according to any one of claims 1-16, wherein the nucleic acid molecules encoding the repertoire of polypeptides are arranged in an ordered fashion to create an array.
19. A method according to claim 18, wherein the array is achieved using robotic transfer.

20. A method according to any preceding claim, wherein the repertoire of polypeptides is screened separately against two or more target molecules.
- 5 21. A naïve polypeptide repertoire wherein the non-functional members have been completely or partially eliminated by binding with one or more generic ligands, said library or repertoire being suitable for direct screening with a specific target ligand.
22. A naïve polypeptide repertoire designed for selection with both generic and target
10 ligands.
23. A naïve polypeptide repertoire according to claims 21 and 22, encoding or comprising polypeptides of the immunoglobulin superfamily.
- 15 25. A naïve polypeptide repertoire according to claim 23 wherein the polypeptides are antibody or T-cell receptor polypeptides.
26. A naïve polypeptide repertoire according to claim 25, wherein the polypeptides are V_H or V_β domains.
20
27. A naïve polypeptide repertoire according to claim 26, wherein the polypeptides are V_L or V_α domains.
28. A naïve polypeptide repertoire wherein a repertoire of polypeptides according to
25 claim 26 and a repertoire of polypeptides according to claim 27 are contacted with generic ligands and the subsets thereby obtained are then pooled.
29. A naïve polypeptide repertoire according to any one of claims 21 to 28, wherein the functional members of the repertoire have a known main-chain conformation.
- 30 30. A naïve polypeptide repertoire according to claim 29, wherein the functional members of the repertoire have a single main-chain conformation.

31. A naïve polypeptide repertoire according to claims 29 or 30, wherein the immunoglobulin scaffold is based on germline V gene segment sequences.

32. A naïve polypeptide repertoire according to any one of claims 21 to 30, wherein
5 the polypeptides are varied at random positions.

33. A naïve polypeptide repertoire according to any one of claims 21 to 30, wherein the polypeptides are varied at selected positions.

10 34. A naïve polypeptide repertoire according to claim 33, wherein the selected positions are those which form the binding site for the target ligand.

35. A naïve polypeptide repertoire according to claim 34, wherein the selected positions are a subset of those which form the binding site for the target ligand.

15 36. A naïve polypeptide repertoire wherein a repertoire of polypeptides according to claim 35 is first contacted with a target ligand in order to isolate a subset of polypeptides specific for the target ligand, the subset of polypeptides then being varied at a further subset of residues in order to modify the function, specificity or affinity of target ligand
20 interaction.

37. A naïve polypeptide repertoire according to claims 33-36, wherein the variation is achieved by incorporating all 20 different amino acids at the positions to be varied.

25 38. A naïve polypeptide repertoire according to claims 33-36, wherein the variation is achieved by incorporating some but not all of the 20 different amino acids at the positions to be varied

39. A nucleic acid library encoding a repertoire of polypeptides according to any one
30 of claims 21 to 38.

40. A repertoire of polypeptides according to any one of claims 21 to 38, which repertoire is immobilised in the form of an array on a solid support.

41. A repertoire according to claim 40, wherein the polypeptides of the repertoire are arrayed by binding to a generic ligand which is immobilised on the solid support.

5 42. A repertoire according to claim 40, wherein the polypeptides are arrayed by direct binding to the solid support.

43. A repertoire of polypeptides according to any one of claims 21 to 38 which is screened in solution form.

Abstract

The invention describes a method for isolating, from a naïve polypeptide repertoire which has not been preselected with a specific target ligand, a polypeptide of interest capable of
5 interacting with said specific target ligand, which method comprises direct screening of the naïve polypeptide repertoire with the target ligand in order to identify the polypeptide of interest.

1/5

Figure 1

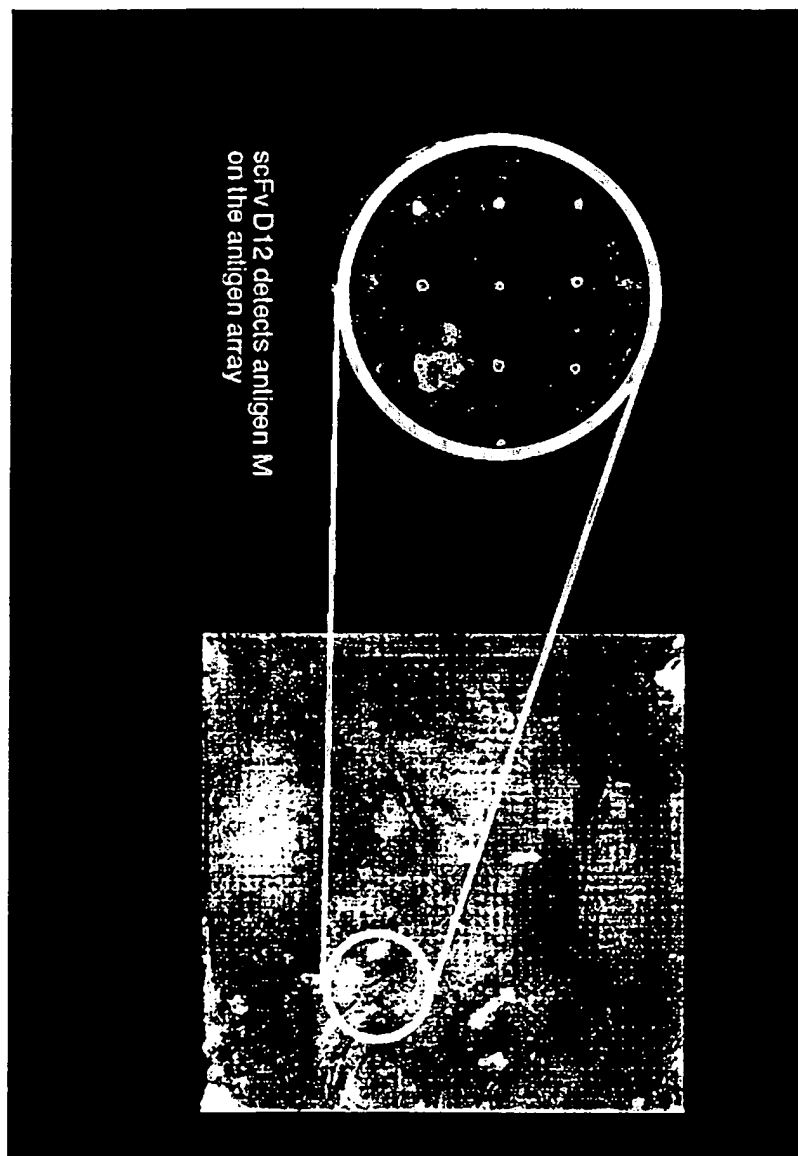
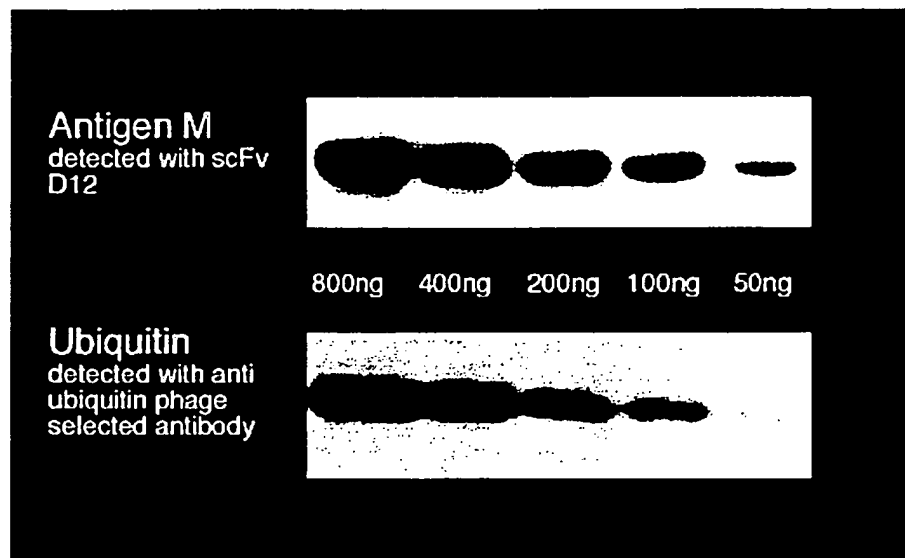


Figure 2

Antigen name	Size	ScFv	CDR H2	CDR H3	CDR L2	CDR L3
M	21.1 kD	D12	GIITSGDAMSYADSVKG	AYTSFDY	NASNLIQS	QQYNATPGT
B	12.0 kD	G12	RIDPTGNTISYADSVKG	VAEIFDY	RASRLQS	QQALSLPTT
C	14.0 kD	H11	SITLAGASTSYADSVKG	ASRSFDY	NASSLQS	QRISPRPTT
O	8.8 kD	C2	GIYASGSTITAYADSVKG	AATSFDY	SASYLQS	QQATTSPNT

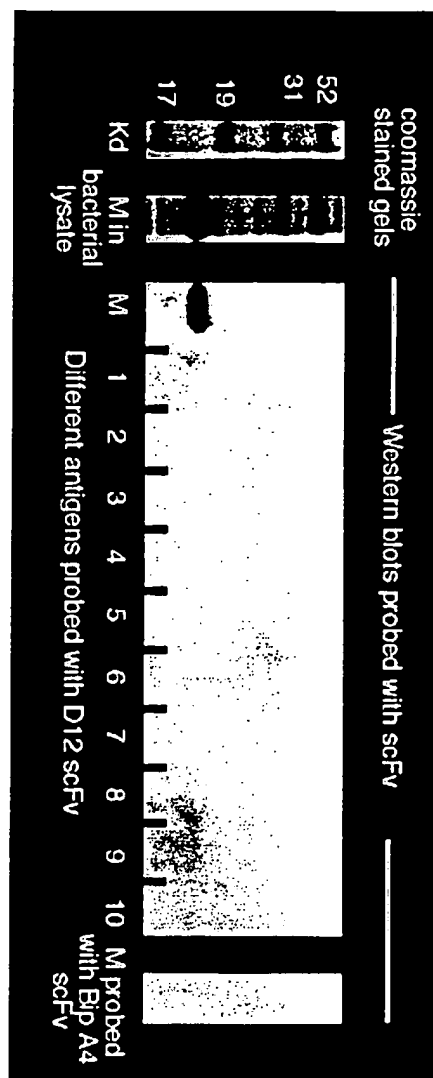
3/5

Figure 3:



4/5

Figure 4:



5/5

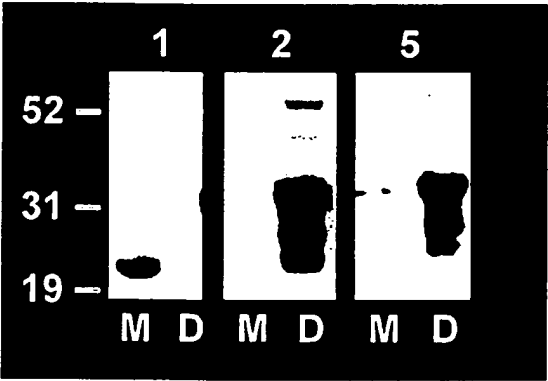


Figure 5

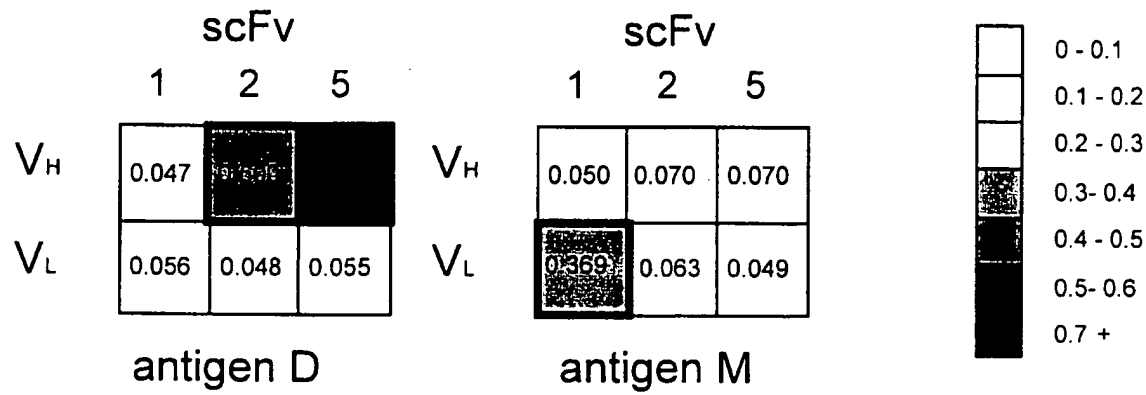


Figure 6

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 June 2001 (07.06.2001)

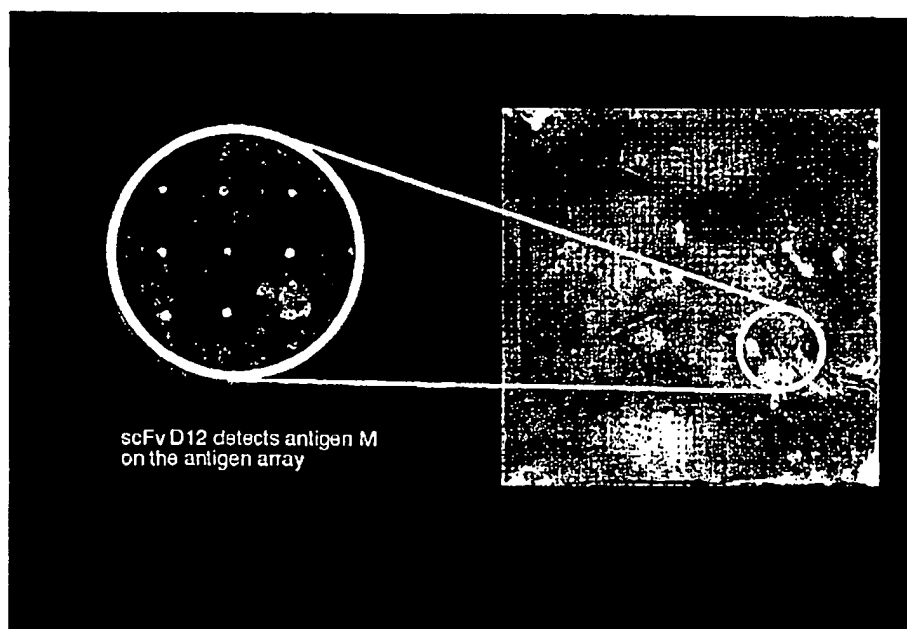
PCT

(10) International Publication Number
WO 01/40312 A3

- (51) International Patent Classification⁷: G01N 33/53, 33/68, C12N 15/10, C07K 1/04 TOMLINSON, Ian [GB/GB]; 13 Tunwells Lane, Great Shelford, Cambridge CB2 5LJ (GB).
- (21) International Application Number: PCT/GB00/04629 (74) Agents: MASCHIO, Antonio et al.; D Young & Co, 21 New Fetter Lane, London EC4A 1AD (GB).
- (22) International Filing Date: 4 December 2000 (04.12.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 9928789.8 3 December 1999 (03.12.1999) GB (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (71) Applicant (*for all designated States except US*): DIVERSYS LIMITED [GB/GB]; 20 Park Crescent, London W1N 4AL (GB).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): HOLT, Lucy, Jessica [GB/GB]; 9b Anstey Way, Trumpington, Cambridge CB2 2JE (GB). DE WILDT, Rudolf, Maria, Theodora [GB/GB]; 27 Grange Road, Cambridge CB3 9AU (GB).
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:
— with international search report

[Continued on next page]

(54) Title: NAIVE SCREENING METHOD



(57) Abstract: The invention describes a method for isolating, from a naïve polypeptide repertoire which has not been preselected with a specific target ligand, a polypeptide of interest capable of interacting with said specific target ligand, which method comprises direct screening of the naïve polypeptide repertoire with the target ligand in order to identify the polypeptide of interest.



(88) Date of publication of the international search report:
17 January 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

In International Application No

PCT/GB 00/04629

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G01N33/53 G01N33/68 C12N15/10 C07K1/04

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 20749 A (MEDICAL RES COUNCIL ;TOMLINSON IAN (GB); WINTER GREG (GB)) 29 April 1999 (1999-04-29) cited in the application abstract page 25, line 11 -page 27, line 2 -----	1-43
A	WO 92 10092 A (AFFYMAX TECH NV) 25 June 1992 (1992-06-25) abstract; claims 17-25 -----	1-43
A	WO 90 15070 A (AFFYMAX TECH NV) 13 December 1990 (1990-12-13) the whole document -----	1-43

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *8* document member of the same patent family

Date of the actual completion of the international search

30 May 2001

Date of mailing of the international search report

13/06/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Gundlach, B

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/GB 00/04629

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9920749 A	29-04-1999	AU 9548798 A EP 1025218 A	10-05-1999 09-08-2000
WO 9210092 A	25-06-1992	AT 199054 T AU 663300 B AU 9153491 A CA 2097708 A DE 69132531 D EP 1046421 A EP 0562025 A JP 6504997 T MX 9102400 A NZ 240744 A US 6124102 A US 5744101 A US 5489678 A US 5889165 A US 5753788 A US 5744305 A US 5770456 A US 5424186 A ZA 9109650 A	15-02-2001 05-10-1995 08-07-1992 07-06-1992 15-03-2001 25-10-2000 29-09-1993 09-06-1994 01-06-1992 27-04-1994 26-09-2000 28-04-1998 06-02-1996 30-03-1999 19-05-1998 28-04-1998 23-06-1998 13-06-1995 07-06-1993
WO 9015070 A	13-12-1990	US 5143854 A AT 110738 T AT 175421 T AU 651795 B AU 5837190 A AU 672723 B AU 7765594 A BR 9007425 A CA 2054706 A DE 69012119 D DE 69012119 T DE 69032888 D DE 69032888 T DK 476014 T DK 619321 T EP 0476014 A EP 0619321 A EP 0902034 A EP 0953835 A ES 2058921 T ES 2129101 T GB 2248840 A, B HK 61395 A HK 64195 A HU 59938 A IL 94551 A JP 11315095 A JP 11021293 A JP 4505763 T KR 9701577 B KR 9701578 B NL 191992 B NL 9022056 T NO 301233 B NZ 233886 A	01-09-1992 15-09-1994 15-01-1999 04-08-1994 07-01-1991 10-10-1996 04-05-1995 21-07-1992 08-12-1990 06-10-1994 22-12-1994 18-02-1999 29-07-1999 14-11-1994 30-08-1999 25-03-1992 12-10-1994 17-03-1999 03-11-1999 01-11-1994 01-06-1999 22-04-1992 05-05-1995 05-05-1995 28-07-1992 30-03-1995 16-11-1999 26-01-1999 08-10-1992 11-02-1997 11-02-1997 01-08-1996 02-03-1992 29-09-1997 25-02-1993

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/GB 00/04629

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9015070 A		SG 13595 G	16-06-1995
		RU 2107072 C	20-03-1998
		US 5925525 A	20-07-1999
		US 6197506 B	06-03-2001
		US 6124102 A	26-09-2000
		US 5744101 A	28-04-1998
		US 5489678 A	06-02-1996
		US 5889165 A	30-03-1999
		US 5753788 A	19-05-1998
		US 6225625 B	01-05-2001
		US 5744305 A	28-04-1998
		US 5547839 A	20-08-1996
		US 5770456 A	23-06-1998